# Degree of Freedom –

Consider counting the population of a school on a sample of 5 classes,
$$[213, 180, 175, 242, 295]$$

Sample Statistics.

| Mean | 221 |
|---|---|
| Median | 213 |
| Std dev | 19.4 |
| Skewness | 0.85 |
| Kurtosis | -0.16 |

So what is the sample mean?
- We used word sample because sample statistics are just estimates of theorised population value.
- This sample are just statistics (of sample) of whole population.

So degree of freedom (DF) are the number of pieces of information we have to estimate population values.

$$Population \xleftarrow{estimates} Sample.$$
$$\mu \qquad\qquad \bar{x}$$
$$\sigma \qquad\qquad S$$

## Degree of Freedom in Descriptive Statistics /ML

Contingency table.

| | Coke | Sprite | Total |
|---|---|---|---|
| Male | | | 20 |
| Female | | | 20 |
| Total | 22 | 18 | 40 |

Degree of freedom = (rows -1) * (columns -1)
$$= (2-1) * (2-1)$$
$$= 1.$$

So, 1 is the df. So I have only 1 cells to vary.

Therefore out of 4 cell, only 1 cell can be varied, remaining 3 we need can predict.

Contigency tabe is ☐ that is of 4 cell

Suppose one cell is 9.
So, only one value has freedom remaining all need to add up.

| 9 ① | 11 ② |
|---|---|
| 13 | 7° |

- More the degree of freedom will be, more the chi-square distribution will look like a normal distribution.

| | Coke | Sprite | Thumbs up | Total |
|---|---|---|---|---|
| Male | | | | 20 |
| Female | | | | 20 |
| Total | 20 | 12 | 8 | 40 |

df = (rows -1) (columns -1)
$$= (2-1)(3-1)$$
$$= (1)(2) = 2.$$

So degree of freedom is 2.
So only 2 cell have freedom to vary, once they are determined other sums off.

Suppose 1 values each is 8

| 8 ① | | |
|---|---|---|
| 12 ② | | |

But we cannot predict other value,

So we need 2 values. Suppose 8 and 6

| 8 ① | 6 ② | 6 ③ |
|---|---|---|
| 12 ② | 6 ④ | 2 ⑤ |

Degree of freedom definition – It is the number of independent values or quantities which can be assigned to a statistics distribution.